SCIENTIFIC PAPERS

Comparison of Surgical Outcomes Between Teaching and Nonteaching Hospitals in the Department of Veterans Affairs

Shukri F. Khuri, MD,*
Samer F. Najjar, MD,*
Jennifer Daley, MD,‡
Barbara Krasnicka, PhD,§
Monir Hossain, MS,‡
William G. Henderson, PhD,§
J. Bradley Aust, MD,‖
Barbara Bass, MD,¶
Michael J. Bishop, MD,#
John Demakis, MD,**
Ralph DePalma, MD,**
Peter J. Fabri, MD,††
Aaron Fink, MD,
James Gibbs, PhD,§
Frederick Grover, MD,§§
Karl Hammermeister, MD,§§
Gerald McDonald, MD,**
Leigh Neumayer, MD,‖‖
Robert H. Roswell, MD,¶¶
Jeannette Spencer, MS, RN,†
Richard H. Turnage, MD,##

From the *VA Boston Healthcare System, West Roxbury, Massachusetts, Harvard Medical School, Boston, Massachusetts, Brigham and Women's Hospital, Boston, Massachusetts; † VA Boston Healthcare System, West Roxbury, Massachusetts; ‡ VA Boston Healthcare System, West Roxbury, Massachusetts, Institute for Health Policy, Massachusetts General Hospital/Partners HealthCare System, Harvard Medical School, Boston, Massachusetts; §Hines VA Cooperative Studies Program Coordinating Center, Hines, Illinois; ‖Department of Surgery, University of Texas Health Science Center, San Antonio, Texas; ¶VA Maryland Health Care System, University of Maryland, Baltimore, Maryland; #National Anesthesia Services, VA Puget Health Care System, Seattle, University of Washington, Seattle, Washington; **Health Services Research and Development Service, Department of Veterans Affairs Headquarters, Washington DC; †† Tampa VA Medical Center, University of South Florida College of Medicine, Tampa, Florida; ‡‡ Atlanta VA Medical Center, Decatur, Georgia, Emory University School of Medicine, Atlanta, Georgia; §§Denver VA Medical Center, University of Colorado Health Sciences Center, Denver, Colorado; ‖‖ VA Salt Lake City Health Care System, University of Utah Medical Center, Salt Lake City, Utah; ¶¶Veterans Integrated Service Network 8, Bay Pines, Florida; ##VA North Texas Health Care System, University of Texas Southwestern Medical School, Dallas, Texas

Presented at the 121st Annual Meeting of the American Surgical Association, April 26-28, 2001, the Broadmoor Hotel, Colorado Springs, Colorado.

Correspondence: Shukri F. Khuri, MD, Surgical Service, VA Boston Healthcare System, 1400 VFW Parkway, West Roxbury, MA 02132.

E-mail: shukri.khuri@med.va.gov

Objective

To determine whether the investment in postgraduate education and training places patients at risk for worse outcomes and higher costs than if medical and surgical care was delivered in nonteaching settings.

Summary Background Data

The Veterans Health Administration (VA) plays a major role in the training of medical students, residents, and fellows.

Methods

The database of the VA National Surgical Quality Improvement Program was analyzed for all major noncardiac operations performed during fiscal years 1997, 1998, and 1999. Teaching status of a hospital was determined on the basis of a background and structure questionnaire that was independently verified by a research fellow. Stepwise logistic regression was used to construct separate models predictive of 30-day mortality and morbidity for each of seven surgical specialties and eight operations. Based on these models, a severity index for each patient was calculated. Hierarchical logistic regression models were then created to examine the relationship between teaching versus nonteaching hospitals and 30-day postoperative mortality and morbidity, after adjusting for patient severity.

Results

Teaching hospitals performed 81% of the total surgical workload and 90% of the major surgery workload. In most specialties in teaching hospitals, the residents were the primary surgeons in more than 90% of the operations. Compared with nonteaching hospitals, the patient populations in teaching hospitals had a higher prevalence of risk factors, underwent more complex operations, and had longer operation times. Risk-adjusted mortality rates were not different between the teaching and nonteaching hospitals in the specialties and operations studied. The unadjusted complication rate was higher in teaching hospitals in six of seven specialties and four of eight operations. Risk adjustment did not eliminate completely these differences, probably reflecting the relatively poor predictive validity of some of the risk adjustment models for morbidity. Length of stay after major operations was not consistently different between teaching and nonteaching hospitals.

Conclusion

Compared with nonteaching hospitals, teaching hospitals in the VA perform the majority of complex and high-risk major procedures, with comparable risk-adjusted 30-day mortality rates. Risk-adjusted 30-day morbidity rates in teaching hospitals are higher in some specialties and operations than in nonteaching hospitals. Although this may reflect the weak predictive validity of some of the risk adjustment models for morbidity, it may also represent suboptimal processes and structures of care that are unique to teaching hospitals. Despite good quality of care in teaching hospitals, as evidenced by the 30-day mortality data, efforts should be made to examine further the structures and processes of surgical care prevailing in these hospitals.

Introduction

Academic health centers and their teaching programs are under duress as they struggle to cope with new and formidable fiscal challenges to their missions of patient care, research, and teaching. [1] [2] [3] [4] [5] [6] Studies conducted in nonfederal healthcare institutions in the United States suggest that patient outcomes in large urban teaching hospitals are better than patient outcomes in smaller community and nonteaching hospitals, [7] [8] [9] [10] although the cost of care is higher [5] [11] [12] [13] [14] [15] [16] and the length of stay is longer [17] in the large teaching hospitals. Most of these studies are based on analyses of administrative databases and examine either all hospital discharges during a finite period or the outcomes of specific, mostly medical, conditions. With the exception of a recent study using the Health Care Financing Administration database of Medicare discharges between 1984 and 1993 for ten medical diagnoses and ten surgical operations, [17] the studies that address the association between teaching and nonteaching hospitals on surgical outcomes have been limited to small series revolving around isolated surgical procedures. [18] [19] [20]

The Veterans Health Administration (VA), the largest single-provider healthcare delivery system in the United States, plays a major role in the training of medical students, residents, and fellows. Approximately one third of postgraduate physicians (residents) and half of all medical students in the United States receive all or part of their clinical training at VA medical treatment facilities each year. Nearly two thirds of all physicians in the United States have received at least some of their training through the VA. [21] The debate about the future of teaching programs in hospitals is also becoming germane to the VA as it struggles to reorganize itself into a more cost-efficient healthcare delivery system. [21] A credible justification for the VA to continue expensive national surgical educational programs may be the demonstration of high quality and value in the VA medical centers (VAMCs) with surgical teaching programs. In this study we compared the quality of surgical care in teaching versus nonteaching VAMCs using data from the VA National Surgical Quality Improvement Program (NSQIP). [22]

METHODS

NSQIP Database

At each of 128 VA surgical centers, a dedicated clinical nurse reviewer prospectively collects preoperative, intraoperative, and 30-day outcome information on all major operations. [22] [23] A major operation is defined as one performed under general, spinal, or epidural anesthesia, in addition to all carotid endarterectomies and inguinal herniorrhaphies regardless of anesthesia type. Data on all surgical specialties, except for cardiac surgery, are transmitted electronically to the VA Cooperative Studies Program Coordinating Center at the VA in Hines, Illinois, where they are cleaned, entered into the database, and analyzed. The data elements contained in the database have been described previously. [23] They include whether the primary surgeon is a resident or a staff surgeon, as well as the postgraduate year (PGY) of the resident if he or she is the primary surgeon. They also include a score of the technical complexity of the operation represented by each CPT-4 code. The complexity score was developed by panels of specialists who graded the complexity of each operation in their specialty on a scale of 1 to 5. Complexity of each operation was assessed based on factors above and beyond the expected patient risk factors that a typical patient having the operation would bring to the operating table. The NSQIP database for three consecutive fiscal years (FY 1997-1999) was used for this study. The outcomes measured were 30-day postoperative mortality, [24] 30-day postoperative morbidity (defined as one or more complications [25]), and postoperative length of hospital stay.

Definition and Identification of Teaching and Nonteaching Hospitals

All VAMCs participating in the NSQIP receive an annual background and structure questionnaire. The survey covers information about the structure of each surgical service and is completed by the NSQIP clinical nurse reviewer with input from the administrative officer of the surgical service. A section of the survey addresses the number of surgical residents assigned to the VAMC in every specialty, and whether the VAMC has a Dean's Committee. The Dean's Committee oversees and addresses issues related to the affiliation between an academic institution and the VAMC.

The initial response rate to the FY 1999 background and structure questionnaire was 90%. A surgical research fellow obtained the information from the remaining 10% of hospitals and verified information from the other hospitals. A hospital was defined as a teaching hospital in a surgical specialty if the VAMC had a Dean's Committee and had at least one resident rotating and operating in that specialty. There was no hospital that had a Dean's Committee without having at least one resident rotating in a surgical specialty. The number of residents rotating in each specialty was not constant throughout the study period from 1997 to 1999. In this study, the number of residents rotating in each specialty is reported on the basis of FY 1999 data.

Statistical Analyses

Statistical analyses were performed separately for each of seven surgical specialties (general, otolaryngology, neurosurgery, vascular, orthopedic, thoracic [noncardiac], urology) and for eight specific operations (carotid endarterectomy, abdominal aortic aneurysm repair, infrainguinal vascular reconstruction, colectomy, open and laparoscopic cholecystectomy, total hip replacement, and lobectomy/pneumonectomy). Hospitals with fewer than 50 cases during the 3-year study period for a given specialty analysis were excluded because the estimates of parameters for those hospitals are unreliable as a result of small sample size. For each surgical specialty or operation, the following steps were taken. First, we compared patient characteristics between the teaching and nonteaching hospitals using chi-square and *t* tests. Second, we identified candidate risk variables to adjust for patient severity from previous work. [24] Third, we imputed missing values using a regression technique that estimates the missing value on the basis of the patient's other nonmissing variables. [26] Fourth, we calculated a severity index for each patient using a logistic regression analysis with mortality (or morbidity) as the dependent variable and patient risk factors as the independent variables. Fifth, we used a hierarchical logistic regression model [27] [28] [29] with mortality (or morbidity) as the dependent variable and only one patient-level independent variable (severity index). Other terms in the model included the intercept indicating interhospital differences in mortality rates, affiliation status (teaching hospital vs. nonteaching hospital), and an interaction term of affiliation by severity. Mean lengths of stay for individual operations between teaching and nonteaching hospitals were tested using the Wilcoxon rank-sum test. All results were considered statistically significant at $P \leq .05$.

RESULTS

Operations

Between October 1, 1996, and September 30, 1999, 690,811 noncardiac operations were performed in 128 VAMCs, of which 433,186 (62.7%) were major operations. Of the 128 hospitals, 95 (74.2%) had a Dean's Committee and at least one surgical resident performing surgery in at least one surgical specialty during FY 1999. These VAMCs were designated as teaching hospitals. The remaining 33 hospitals (25.8%) were designated as nonteaching hospitals. Teaching hospitals performed 80% of the total operations; 57% of the operations were major. In contrast, nonteaching hospitals performed 20% of all operations, and only 26% of the operations were major.

The distribution of major surgery volume in the teaching and nonteaching hospitals is shown in Table 1 by surgical specialty. These data form the basis of the analysis file for this study. They do not include all the major operations in the database because of the exclusion of operations performed within 30 days of another major operation, and the exclusion of hospitals with fewer than 50 operations performed in a given surgical specialty during the 3 years of the study. In both teaching and nonteaching hospitals, the most frequently performed operations were in general surgery, followed by orthopedics and urology. Teaching versus nonteaching hospital status varied widely among specialties. Nonteaching hospitals constituted 37% of the hospitals performing urologic surgery and only 10% of the hospitals performing neurosurgery. For each specialty, the teaching hospitals had significantly higher volumes of major operations than the nonteaching hospitals.

Table 1. DISTRIBUTION OF MAJOR OPERATIONS (1997-1999)

	Teaching Hospitals			Nonteaching Hospitals		
	Number of Operations/Hospital			**Number of Operations/Hospital**		
No. of Hospitals (%)						

	†[a]	Mean ± SD	Range	No. of Hospitals (%)	Mean ± SD	Range	P Value[*]
General surgery	92 (73)	852 ± 253	116 -1,528	34 (27)	427 ± 210	84 - 767	<.0001
Orthopedic surgery	75 (70)	640 ± 230	284 -1,325	32 (30)	338 ± 190	83 - 791	.0001
Urology	75 (63)	427 ± 135	71 -996	45 (37)	188 ± 115	55 - 602	.0001
Vascular surgery	82 (89)	358 ± 147	91 -767	10 (11)	156 ± 76	52 - 278	.0001
Otolaryngology	68 (86)	211 ± 76	52 -517	11 (14)	110 ± 73	51 - 303	.0009
Neurosurgery	51 (90)	316 ± 167	68 -1,112	6 (10)	175 ± 77	93 - 301	.0094
Thoracic surgery	66 (87)	142 ± 69	53 -436	10 (13)	98 ± 45	52 - 200	.0228

* Difference between mean number of operations performed in teaching and nonteaching hospitals.

† Hospitals with fewer than 50 operations in a specialty between 1997 and 1999 were excluded from the analysis.

In the teaching hospitals, the resident was the primary surgeon in 84% to 95% of the cases (Table 2). The distribution of the PGY level of the primary surgeon in the seven surgical specialties is shown in Figure 1. Most of the operations were performed by PGY 5 residents. Although by definition one would not expect residents to be operating in nonteaching hospitals, Table 2 shows that in a few operations performed in nonteaching hospitals, a resident was the primary surgeon of record. There are two possible reasons for this discrepancy. First, the nonteaching status of a hospital in this study was determined on the basis of the hospital structure in FY 1999; some surgical specialties may have had residents operating in FY 1997 and/or 1998, but not in 1999. Second, the designation of teaching versus nonteaching hospital was made on the basis of residents rotating to a specific specialty. Operations in a designated specialty, such as noncardiac thoracic surgery, may have been performed by residents from another specialty, such as general surgery, in a hospital designated as a nonteaching hospital in thoracic surgery because it did not have thoracic surgery residents rotating through it.

Table 2. SELECTED CHARACTERISTICS OF OPERATIONS

Surgical specialty	Characteristics	Teaching Hospitals	Nonteaching Hospitals	P Value
General surgery	% of operations performed by resident	91.2	1.4	<.0001
	Mean operation time (hr)	1.9 ± 1.5	1.3 ± 1.0	<.0001
	Complexity score[*] (mean ± SD)	2.43 ± 0.62	2.36 ± 0.56	.0001
Orthopedic surgery	% of operations performed by resident	92.6	1.3	<.0001
	Mean operation time (hr)	2.0 ± 1.2	1.3 ± 1.0	<.0001
	Complexity score (mean ± SD)	3.08 ± 0.77	2.89 ± 0.72	<.0001
Urology	% of operations performed by resident	95.3	5.3	<.0001
	Mean operation time (hr)	1.8 ± 1.6	1.1 ± 1.1	<.0001
	Complexity score (mean ± SD)	2.5 ± 0.90	2.30 ± 0.71	<.0001
Vascular surgery	% of operations performed by resident	87.9	15.3	.001
	Mean operation time (hr)	3.2 ± 1.9	2.8 ± 1.8	.018
	Complexity score (mean ± SD)	3.54 ± 1.07	3.39 ± 1.14	.0001
Otolaryngology	% of operations performed by resident	95.6	13.5	<.0001
	Mean operation time (hr)	2.9 ± 2.8	1.7 ± 1.9	<.0001
	Complexity score (mean ± SD)	2.25 ± 0.65	2.21 ± 0.65	.021
Neurosurgery	% of operations performed by resident	84.4	2.4	.001
	Mean operation time (hr)	3.2 ± 1.8	2.6 ± 1.3	.0196
	Complexity score (mean ± SD)	2.64 ± 0.74	2.45 ± 0.60	.0001
Thoracic surgery	% of operations performed by resident	85.6	14.6	<.0001
	Mean operation time (hr)	2.3 ± 1.9	2.0 ± 1.7	.0535
	Complexity score (mean ± SD)	2.33 ± 0.66	2.37 ± 0.66	.0493

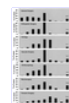* Complexity score ranges from 1-5, with 1 = least complex and 5 = most complex.

Figure 1. Distribution of the postgraduate year (PGY) level of the primary surgeon in seven surgical specialties in 95 teaching VA medical centers. SS, staff surgeon. The percentage of operations in each specialty performed by an SS is shown numerically over the corresponding barograph.

Table 2 also shows the mean operating times and complexity scores for all operations in each specialty in teaching and nonteaching hospitals. Both the mean operating time and the mean complexity score were significantly higher in the teaching hospitals across all specialties (except for the complexity score in noncardiac thoracic surgery).

Preoperative Patient Characteristics

Table 3 compares important preoperative patient characteristics in the teaching and nonteaching hospitals. In many of the surgical specialties, more patients in teaching hospitals tended to be nonwhite; patients in teaching hospitals had a higher percentage of functional impairment, weight loss, steroid use, and emergency operations. Patients in teaching hospitals had significantly lower preoperative serum albumin levels and were more likely to have a low hematocrit and elevated serum bilirubin, serum aspartate aminotransferase (SGOT), and blood urea nitrogen.

Table 3. PREOPERATIVE PATIENT CHARACTERISTICS

Preoperative Patient Characteristics	General			Orthopedic			Urology			Vascular		
	TH	NTH	P	TH	NTH	P	TH	NTH	P	TH	NTH	P
Mean age (yr)	60.8	61.8	.0001	58.0	56.8	.0001	65.2	67.2	<.0001	66.1	66.7	.0286
White (%)	69.1	83.2	.0001	69.9	81.6	.001	70.0	86.6	<.0001	77.5	86.7	<.0001
Functional status >0 (%)[*]	12.1	10.5	<.0001	15.7	10.5	<.0001	7.7	9.0	.0001	19.2	13.4	<.0001

Preoperative Patient Characteristics	TH	NTH	P	TH	NTH	P	TH	NTH	P	TH	NTH	P
>10% weight loss in 6 mo	6.0	3.5	<.0001	1.3	0.9	.001	1.7	1.2	.0003	2.9	2.6	.3653
Steroid use (%)	3.1	2.5	<.0001	2.9	2.0	.001	2.4	2.3	.363	2.6	2.2	.2397
Impaired sensorium (%)	2.7	3.0	.0589	2.5	3.2	<.0001	1.1	2.1	<.0001	2.8	2.2	.1325
Emergency (%)	13.3	10.8	<.0001	5.6	5.6	.976	2.3	1.6	<.0001	8.7	6.3	<.0001
Hematocrit $\leq$38% (%)	30.0	19.7	<.0001	24.4	16.2	.001	24.6	21.8	<.0001	43.1	27.8	<.0001
Serum albumin $\leq$3.5 mg/dL (%)	26.4	22.4	.0001	15.3	13.0	.001	12.2	13.6	.001	31.2	23.5	.0001
Serum bilirubin >1.0 mg/dL (%)	16.6	11.9	<.0001	6.3	6.3	.9659	4.3	5.2	.0008	5.4	5.3	.8128
SGOT >40 IU/ml (%)	17.4	11.7	<.0001	10.2	9.4	.008	4.4	4.8	.0845	9.6	8.8	.2184
Blood urea nitrogen >40 mg/dL (%)	2.5	4.0	<.0001	2.0	1.2	.001	2.6	3.0	.0669	7.2	4.4	<.0001
Blood urea nitrogen <40 mg/dL (%)	60.8	NTH	<.0001	70.4	NTH	<.0001	70.0	NTH	<.0001	77.5	NTH	<.0001

Preoperative Patient Characteristics	Otolaryngology			Neurosurgery			Noncardiac Thoracic		
	TH	NTH	P	TH	NTH	P	TH	NTH	P
Mean age (yrs)	57.0	56.8	.5662	55.1	54.6	.2179	63.2	64.2	.0074
Caucasian (%)	69.8	80.9	<.0001	70.4	87.0	.001	76.7	86.7	<.0001
Functional status >0 (%)	9.2	6.2	.0013	13.8	13.3	.631	12.4	10.4	.059
>10% weight loss in months	6.0	3.5	.0002	2.2	1.1	.018	11.3	7.0	<.0001
Steroid use (%)	3.9	2.9	.0691	4.1	2.8	.032	5.0	3.2	.0099
Impaired sensorium (%)	2.6	3.0	.4541	6.3	4.1	.004	3.4	3.4	.9999
Emergency (%)	2.4	1.7	.1263	6.1	4.0	.005	6.4	4.8	.0542
Hematocrit 38% (%)	21.7	14.0	<.0001	16.3	14.9	.247	42.7	30.6	<.0001
Serum albumin 3.5 mg/dL (%)	14.9	10.6	.0001	9.6	8.4	.1891	37.5	30.9	.0001
Serum bilirubin >1.0 mg/dL (%)	6.5	5.2	.0719	3.5	2.6	.105	8.5	8.0	.5494
SGOT >40 IU/ml (%)	10.6	8.5	.0274	7.8	7.2	.473	16.4	16.4	.9715
Blood urea nitrogen >40 mg/dL (%)	2.6	1.4	.01	1.0	0.8	.4386	4.6	3.4	.0915

TH, teaching hospital; NTH, nonteaching hospital; SGOT, blood aspartate aminotransferase.

* Functional status: 0, independent in activities of daily living; 1, partially dependent; 2, totally dependent.

Thirty-Day Mortality Rates

Table 4 presents the unadjusted 30-day mortality rates for each specialty in teaching and nonteaching hospitals. In all specialties, the mortality rate in teaching hospitals was higher than in nonteaching hospitals, with the difference reaching statistical significance in general surgery and orthopedic surgery.

Table 4. UNADJUSTED 30-DAY MORTALITY AND MORBIDITY RATES

Specialty	30-Day Mortality Rate (%)			30-Day Morbidity Rate (%)		
	TH	NTH	P	TH	NTH	P
General surgery	3.6	2.8	.001	13.0	7.3	<.0001
Orthopedic surgery	1.6	1.1	.001	6.5	4.5	<.0001
Urology	0.8	0.8	.453	6.0	3.4	<.0001
Vascular surgery	3.9	3.2	.172	17.3	13.3	<.0001
Otolaryngology	2.5	1.7	.087	7.8	5.4	.003
Neurosurgery	2.3	1.7	.193	9.0	6.5	.005
Thoracic surgery	5.7	4.8	.267	16.6	14.8	.136

TH, teaching hospitals; NTH, nonteaching hospitals.

To risk-adjust the mortality rates, models predictive of 30-day mortality in the FY 1997 to 1999 database were developed for each specialty using stepwise logistic regression. The models had high predictive validity, as evidenced by their respective c-indexes: general surgery 0.91, orthopedics 0.92, urology 0.875, vascular surgery 0.82, otolaryngology 0.93, neurosurgery 0.89, and noncardiac thoracic surgery 0.75. The top 10 independent predictors of 30-day mortality for each specialty model were similar to those published previously by the VA National Surgical Risk Study. [24] Based on the predictive models, a severity index was calculated for each patient and was entered as an independent variable along with the teaching status of the hospital in a hierarchical logistic regression model with mortality as the dependent variable. The results of the hierarchical logistic regression analysis are shown in Table 5. The probability value for the severity index was highly significant in each specialty model, indicating that patient severity of illness was a strong predictor of 30-day postoperative mortality. Except for orthopedic surgery, the lack of significance in the probability values related to the teaching status of the hospitals indicates that after adjusting for differences in patient severity between teaching and nonteaching hospitals, the teaching status of the hospital was no longer significantly related to the 30-day postoperative mortality rate.

Table 5. TEACHING STATUS AND SEVERITY INDEX VERSUS MORTALITY AND MORBIDITY RATES BY SPECIALTY

Specialty	30-Day Mortality Rate			30-Day Morbidity Rate		
	Estimate	95% CI	P	Estimate	95% CI	P

General surgery						
Intercept	-4.83	(-5.04, -4.62)	<.0001	-3.04	(-3.22, -2.86)	<.0001
Severity effect	1.02	(0.96, 1.07)	<.0001	1.04	(0.96, 1.11)	<.0001
Teaching effect	0.11	(-0.12, 0.34)	.3433	0.67	(0.46, 0.87)	<.0001
Orthopedic surgery						
Intercept	-6.13	(-6.46, -5.79)	<.0001	-3.60	(-3.72, -3.30)	<.0001
Severity effect	1.05	(0.96, 1.13)	<.0001	1.04	(0.95, 1.17)	<.0001
Teaching effect	0.39	(0.03, 0.75)	.0327	0.45	(0.09, 0.57)	.001
Urology						
Intercept	-5.88	(-6.25, -5.5)	<.0001	-3.63	(-3.94, -3.33)	<.0001
Severity effect	0.99	(0.85, 1.13)	<.0001	1.02	(0.81, 1.23)	<.0001
Teaching effect	0.09	(-0.33, 0.52)	.666	0.48	(0.14, 0.83)	.0054
Vascular surgery						
Intercept	-4.22	(-4.71, -3.73)	<.0001	-2.05	(-2.34, -1.78)	<.0001
Severity effect	1.18	(0.94, 1.43)	<.0001	1.25	(0.97, 1.53)	<.0001
Teaching effect	0.34	(-0.16, 0.84)	.1821	0.38	(0.09, 0.68)	.0103
Otolaryngology						
Intercept	-5.61	(-6.43, -4.79)	<.0001	-3.27	(-3.74, -2.81)	<.0001
Severity effect	1.13	(0.85, 1.4)	<.0001	1.16	(0.79, 1.53)	<.0001
Teaching effect	0.33	(-0.5, 1.17)	.4344	0.27	(-0.22, 0.75)	.2782
Neurosurgery						
Intercept	-5.79	(-6.99, -4.59)	<.0001	-2.84	(-3.24, -2.44)	<.0001
Severity effect	1.42	(0.99, 1.85)	<.0001	0.98	(-0.66, 1.31)	<.0001
Teaching effect	0.96	(-0.23, 2.15)	.1151	0.18	(-0.23, 0.59)	.39
Thoracic surgery						
Intercept	-3.49	(-3.96, -3.02)	<.0001	-1.82	(-2.12, -1.52)	<.0001
Severity effect	1.23	(0.93, 1.54)	<.0001	0.94	(0.59, 1.29)	<.0001
Teaching effect	0.32	(-0.17, 0.8)	.2014	0.06	(-0.25, 0.38)	.6915

95% CI, 95% confidence interval.

Thirty-Day Morbidity Rates

Table 4 also presents the unadjusted 30-day morbidity rates for each specialty in teaching and nonteaching hospitals. In all specialties, except for noncardiac thoracic surgery, the morbidity rate in teaching hospitals was significantly higher than in nonteaching hospitals. The specialty-specific models, based on the FY 1997 to 1999 database, developed to identify predictors of 30-day morbidity, had a lower predictive validity than the models developed for risk adjustment of 30-day mortality. The c-indexes for the morbidity models were general surgery 0.79, orthopedics 0.76, urology 0.70, vascular 0.63, otolaryngology 0.78, neurosurgery 0.73, and noncardiac thoracic 0.67. The top 10 independent predictors of 30-day morbidity for each specialty model were also similar to those published previously by the VA National Surgical Risk Study. [25] The results of the hierarchical logistic regression analysis to determine the relative effects of patient severity and hospital type on the morbidity rate are shown in Table 5. As in the mortality analysis, the probability value for the effect of severity of illness was highly significant in each specialty model, indicating that severity of illness is highly predictive of 30-day postoperative morbidity. Unlike the mortality analysis, however, risk adjustment did not eliminate the effect of hospital type on the observed differences in morbidity rates between teaching and nonteaching hospitals in general surgery, orthopedics, urology, and vascular surgery. In these four surgical specialties, the increased morbidity rates in the teaching hospitals could not be completely explained by the increased severity of illness of the patient populations. Table 6 shows the frequency distribution of the types of complications encountered in each of the seven specialties. Wound complications (including superficial wound infection, deep wound infection, and wound disruption) were the most frequently encountered morbidity in most of the specialties.

Table 6. ADVERSE EVENT RATES

Specialty	Type of Adverse Event	Rate of Adverse Events (%)			% of Total Adverse Event in Specialty	
		TH	NTH	*P*	TH	NTH
General surgery	Wound	4.9	2.7	<.0001	27.1	28.6
	Respiratory	4.9	2.6	<.0001	27.1	28.2
	Urinary	2.7	1.0	<.0001	15.2	10.9
	CNS	0.3	0.3	.2111	1.8	2.7
	Cardiac	1.2	0.7	<.0001	6.8	7.4
	Other	3.9	2.1	<.0001	22.0	22.2
	Total				100	100
Orthopedic surgery	Wound	1.8	1.2	<.0001	23.1	22.1
	Respiratory	1.8	1.3	<.0001	22.7	23.4
	Urinary	1.9	1.4	.0013	23.6	26.3
	CNS	0.3	0.2	.0199	3.6	3.0
	Cardiac	0.6	0.4	.0041	7.8	7.1
	Other	1.5	1.0	<.0001	19.2	18.1
	Total				100	100
Urology	Wound	1.5	0.6	.0001	19.7	15.7
	Respiratory	1.3	0.7	.0005	17.2	20.1
	Urinary	2.6	1.5	.0001	34.3	37.8
	CNS	0.2	0.1	.0869	3.0	3.2
	Cardiac	0.5	0.4	.3014	6.1	9.0
	Other	1.5	0.6	.0001	19.7	14.2
	Total				100	100
Vascular surgery	Wound	5.7	2.9	<.0001	24.9	18.7
	Respiratory	5.4	3.6	.0022	23.5	23.2
	Urinary	3.5	2.1	.0025	15.2	13.3
	CNS	1.2	1.7	.1218	5.3	10.8

Specialty	Event					
	Cardiac	2.4	2.1	.4413	10.6	13.7
	Other	4.7	3.2	.0045	20.5	20.3
	Total				100	100
Otolaryngology	Wound	2.4	3.0	.1642	23.5	52.1
	Respiratory	3.9	1.9	.0003	38.7	32.5
	Urinary	1.3	0.4	.0057	13.0	7.0
	CNS	0.4	1.1	.0874	3.9	1.4
	Cardiac	0.8	0.2	.0119	7.4	2.8
	Other	1.4	0.3	.0013	13.5	4.2
	Total				100	100
Neurosurgery	Wound	2.4	2.8	.4245	20.2	32.2
	Respiratory	3.2	2.3	.1044	27.2	26.7
	Urinary	2.8	1.3	.0014	24.2	15.5
	CNS	1.0	0.3	.0142	8.8	3.3
	Cardiac	0.9	0.6	.3255	7.3	6.7
	Other	1.5	1.3	.7639	12.3	15.6
	Total				100	100
Thoracic surgery	Wound	2.2	2.2	.9472	9.3	10.0
	Respiratory	11.4	10.1	.2256	48.7	47.4
	Urinary	3.2	3.1	.8154	13.8	14.4
	CNS	0.6	0.8	.3785	2.5	3.8
	Cardiac	2.1	2.5	.4582	8.9	11.5
	Other	3.9	2.8	.069	16.8	12.9
	Total				100	100

Reference group is non-teaching hospitals

Adverse Events: Wound = superficial wound infection, deep wound infection, wound disruption

Respiratory = pneumonia, unplanned reintubation for respiratory/cardiac failure, pulmonary embolism, on ventilator for >48 hrs

Urinary = progressive renal insufficiency, acute renal fialure, urinary tract infection

CNS (Central Nervous System) = stroke/CVA (cerebral vascular accident), coma >24 hrs, peripheral nerve injury

Cardiac arrest with CPR (cardiopulmonary resuscitation), myocardial infarction

Other = prolonged ileus/bowel obstruction, bleeding with >4 units RBCs transfused, graft/prosthesis/flap failure, deep vein thrombosis/thrombo phlebitis, systemic sepsis

Mortality and Morbidity Rates in Prevalent Operations

The above analyses, performed on the basis of surgical specialties, were repeated for eight selected operations, based on the CPT-4 code groupings shown in ~~Table 7~~. ~~Table 8~~ lists the operative characteristics and the unadjusted 30-day mortality and morbidity rates for each operation in both types of hospitals. As was observed in the specialty analysis, operations performed in teaching hospitals were mostly performed by residents, had higher complexity scores, and had longer operative times than operations of the same type performed in nonteaching hospitals. Unadjusted mortality rates were not significantly different between teaching and nonteaching hospitals in all eight operation types. In contrast, complication rates were significantly higher in teaching hospitals in four operation types (infrainguinal vascular reconstruction, colectomy, and open and laparoscopic cholecystectomy).

Table 7. CPT-4 CODES USED

Procedure	Codes
Colectomy	44140, 44141, 44143-44147, 44150-44153, 44155, 44156, 44160
Open cholecystectomy	47600, 47605, 47610
Laparoscopic cholecystectomy	56340, 56341, 56342, 49310, 49311
Total hip replacement	27130, 27131, 27132, 27134
Lobectomy/pneumonectomy	32480, 32485, 32490, 32501, 32440, 32445
Abdominal aortic aneurysmectomy	35081
Carotid endarterectomy	35301
Infrainguinal vascular reconstruction	35521, 35533, 35546, 35548, 35549, 35551, 35556, 35558, 35565, 35566, 35571, 35582, 35583, 35587, 35621, 35623, 35646, 35651, 35654, 35656, 35661, 35665, 35666, 35671

Table 8. OPERATIVE CHARACTERISTICS AND MORTALITY AND MORBIDITY RATES

Operation	Number of Operations	Mean Complexity Score	Performed by Resident (%)	Mean Operative Time (hr)	Unadjusted Mortality Rate (%)	Unadjusted Morbidity Rate (%)
Carotid endarterectomy						
TH	7,613	4.24	86.21	2.65 ± 0.9	0.97	6.4
NTH	543	3.74	12.34	2.11 ± 0.9	1.47	5.7
P		.0001	.0001	.0023	.26	.50
Abdominal aortic aneurysmectomy						
TH	2,038	4.15	88.37	4.10 ± 1.6	4.47	26.4
NTH	116	3.77	10.34	3.68 ± 1.7	3.45	31.0
P		.0002	<.0001	.137	.60	.001
Infrainguinal vascular reconstruction						
TH	7,209	3.74	87.06	4.61 ± 1.8	2.83	23.2
NTH	376	3.44	16.22	4.09 ± 2.2	3.46	15.7
P		.0001	<.0001	.058	.48	.0008

Colectomy						
TH	8,038	3.00	93.64	3.17 ± 1.5	6.53	28.97
NTH	1,201	2.95	1.42	2.52 ± 1.3	5.33	21.23
P		.015	<.0001	<.0001	.11	.001
Open cholecystectomy						
TH	2,911	2.73	92.55	2.31 ± 1.2	2.95	17.1
NTH	653	2.78	1.38	1.73 ± 0.9	2.91	10.11
P		.015	<.0001	<.0001	.95	.001
Laparoscopic cholecystectomy						
TH	6,141	2.42	92.30	1.90 ± 0.9	0.60	5.44
NTH	1,411	2.42	1.28	1.55 ± 0.8	0.35	2.76
P		.43	<.0001	<.0001	.26	.001
Total hip replacement						
TH	4,825	4.01	93.60	2.85 ± 1.1	0.87	8.9
NTH	763	3.95	0.92	2.56 ± 1.2	1.31	7.9
P		.002	<.0001	.036	.24	.36
Lobectomy/pneumonectomy						
TH	2,045	2.63	88.41	3.29 ± 1.2	6.06	21.03
NTH	215	2.60	19.07	3.05 ± 1.6	5.12	23.72
P		.66	<.0001	.33	.58	.36

TH, teaching hospital; NTH, nonteaching hospital.

The c-indexes of the predictive models for the eight operations were consistently higher in mortalities than in morbidities: carotid endarterectomy, mortality 0.69, morbidity 0.59; abdominal aortic aneurysm repair, mortality 0.74, morbidity 0.66; infrainguinal vascular reconstruction, mortality 0.77, morbidity 0.59; colectomy, mortality 0.84, morbidity 0.67; open cholecystectomy, mortality 0.84, morbidity 0.69; laparoscopic cholecystectomy, mortality 0.85, morbidity 0.70; total hip replacement, mortality 0.83, morbidity 0.65; and lobectomy/pneumonectomy, mortality 0.69, morbidity 0.63.

Results of the hierarchical modeling are shown in Table 9. None of the operations showed a significant association between hospital type and mortality. For six of the operations (infrainguinal reconstruction, colectomy, open and laparoscopic cholecystectomy, total hip replacement, and lobectomy/pneumonectomy), severity of illness showed a significant association with mortality.

Table 9. TEACHING STATUS AND SEVERITY INDEX VERSUS DEATH AND COMPLICATIONS RATES BY OPERATION

	30-Day Death Rate			30-Day Complications Rate		
Operation	**Estimate**	**95% CI**	***P***	**Estimate**	**95% CI**	***P***
Carotid endarterectomy						
Intercept	-4.29	(-5.07, -3.52)	<.0001	-2.82	(-3.25, -2.39)	<.0001
Severity effect	0.57	(-0.36, 1.49)	.2259	-0.02	(-1.26, 1.22)	.9752
Teaching effect	-0.65	(-1.46, 0.16)	.1182	0.06	(-0.37, 0.50)	.7726
Abdominal aortic aneurysmectomy						
Intercept	-3.60	(-4.78, -2.41)	<.0001	-0.89	(-1.46, -0.31)	.0029
Severity effect	0.78	(-0.24, 1.80)	.1331	1.43	(0.50, 2.37)	.0029
Teaching effect	0.05	(-1.15, 1.25)	.9346	-0.21	(-0.80, 0.37)	.4733
Infrainguinal vascular reconstruction						
Intercept	-4.19	(-5.07, -3.32)	<.0001	-1.75	(-2.13, -1.37)	<.0001
Severity effect	1.48	(0.85, 2.10)	<.0001	1.25	(0.43, 2.08)	.0033
Teaching effect	0.20	(-0.69, 1.09)	.6660	0.53	(0.15, 0.91)	.0069
Colectomy						
Intercept	-3.63	(-4.05, -3.22)	<.0001	-1.45	(-1.67, -1.22)	<.0001
Severity effect	1.00	(0.81, 1.19)	<.0001	0.93	(0.68, 1.18)	<.0001
Teaching effect	0.16	(-0.29, 0.60)	.4826	0.44	(0.19, 0.68)	.0006
Open cholecystectomy						
Intercept	-4.41	(-5.13, -3.68)	<.0001	-2.32	(-2.65, -1.99)	<.0001
Severity effect	0.85	(0.48, 1.23)	<.0001	0.96	(0.55, 1.38)	<.0001
Teaching effect	-0.24	(-1.06, 0.58)	.5618	0.58	(0.23, 0.94)	.0014
Laparoscopic cholecystectomy						
Intercept	-9.30	(-11.67, -6.93)	<.0001	-3.90	(-4.31, -3.48)	<.0001
Severity effect	1.41	(0.56, 2.25)	.0013	-1.14	(-1.58, -0.69)	<.0001
Teaching effect	0.76	(-1.79, 3.31)	.5592	0.82	(0.37, 1.26)	.0003
Total hip replacement						
Intercept	-6.22	(-7.49, -4.95)	<.0001	-2.62	(-3.00, -2.24)	<.0001
Severity effect	1.05	(0.48, 1.61)	.0004	0.81	(0.29, 1.32)	.0024
Teaching effect	-0.39	(-1.78, 1.01)	.5876	0.09	(-0.32, 0.50)	.6704
Lobectomy/pneumonectomy						
Intercept	-3.14	(-3.89, -2.40)	<.0001	-1.29	(-1.76, -0.82)	<.0001
Severity effect	0.92	(0.05, 1.80)	.0388	0.84	(0.19, 1.50)	.0118
Teaching effect	0.13	(-0.64, 0.89)	.7481	-0.14	(-0.63, 0.35)	.5723

95% CI, 95% confidence interval.

In the hierarchical logistic regression modeling for morbidity (see ~~Table 9~~), for all the operations, except for carotid endarterectomy, severity of illness was a predictor of 30-day postoperative morbidity. In four of the eight operations (infrainguinal vascular reconstruction, colectomy, open and closed cholecystectomy), hospital type continued to be a significant predictor of 30-day

postoperative morbidity, after adjustment for patient severity.

Length of Stay

Table 10 shows the mean and median postoperative length of stay in teaching and nonteaching hospitals in eight prevalent operations. The difference in length of stay between the teaching and nonteaching hospitals was not consistent across these operations. It was not significant after open and laparoscopic cholecystectomy. Length of stay was longer in teaching hospitals after carotid endarterectomy, infrainguinal vascular reconstruction, and colectomy; it was shorter after abdominal aortic aneurysmectomy, total hip replacement, and lobectomy/pneumonectomy.

Table 10. POSTOPERATIVE LENGTH OF STAY

Operation	Length of Stay (days)		
	Mean ± SD	Median	*P*[*]
Carotid endarterectomy			
TH	3.1 ± 4.5	2	.0034
NTH	2.7 ± 2.9	2	
Abdominal aortic aneurysmectomy			
TH	11.0 ± 12.6	7	.0003
NTH	12.1 ± 9.4	8	
Infrainguinal vascular reconstruction			
TH	10.0 ± 11.8	7	.0416
NTH	9.4 ± 11.8	6	
Colectomy			
TH	12.1 ± 14.5	8	<.0001
NTH	11.2 ± 9.5	8	
Open cholecystectomy			
TH	7.6 ± 9.3	5	.081
NTH	7.1 ± 6.6	6	
Laparoscopic cholecystectomy			
TH	3.2 ± 5.3	2	.8045
NTH	3.2 ± 5.4	2	
Total hip replacement			
TH	7.4 ± 5.3	6	.0001
NTH	8.0 ± 5.6	7	
Lobectomy/pneumonectomy			
TH	12.0 ± 11.9	8	.0180
NTH	13.1 ± 12.3	9	

TH, teaching hospital; NTH, nonteaching hospital.

* Based on a Wilcoxon rank-sum test between mean length of postoperative stay in teaching and nonteaching VA hospitals.

DISCUSSION

The VA has made a major commitment to the training of healthcare providers. In 1997, the VA funded nearly 9,000 physician residency positions at 131 medical school-affiliated VAMCs, through which more than 34,000 residents rotated annually, and approximately 45,000 positions among other health professions. [30] Surgical training programs today rely increasingly on opportunities for residents to learn and gain surgical experience in the VA as pressures on efficiency and reducing costs in the nonfederal sector increase. In 1999, 75% of the 128 surgical services in the VA were affiliated with medical schools and had surgical teaching programs in one or more surgical specialties. Since 1992, a strict residency supervision policy has been applied in the VA in which the attending surgeon is required to be present or assisting in the operating room on every major operation.

Despite the apparent value and necessity of teaching the next generation of surgeons, some patients and policymakers question whether the investment in postgraduate education and training places patients at risk for worse outcomes and higher costs than if medical and surgical care were delivered in nonteaching settings. In this study, we compared the outcomes of major surgery in teaching and nonteaching surgical programs in the VA. We used prospectively collected data from the NSQIP to adjust for differences in patient preoperative characteristics and the complexity of the operations. The risk adjustment models developed and used by the NSQIP have been validated as measures of quality of surgical care. [22][31][32]

The unadjusted mortality rate was not significantly different between teaching and nonteaching hospitals in five of the seven specialties and in all eight operations studied. After adjustment for patient preoperative characteristics and the complexity of the operations, there were no statistical differences in the adjusted mortality rates in any of the surgical specialties or common procedures (except for orthopedic surgery, where a difference between teaching and nonteaching hospitals was of borderline statistical significance; see Table 5). The lack of stability of the orthopedic model may account for this apparent significance, particularly because no difference in the adjusted mortality rate of total hip replacement was found. Surgeons in VA teaching hospitals operated on sicker and more complex patients than those in nonteaching hospitals with a comparable mortality rate, indicating that the teaching status of a VAMC does not adversely affect 30-day postoperative mortality rate.

The NSQIP also collects information on the incidence of 20 postoperative adverse events and defines the presence of postoperative morbidity as the occurrence of any one or more of these adverse events. The unadjusted morbidity rates were significantly lower in six of the seven surgical specialties and in five of the eight common major operations in patients operated on in nonteaching hospitals. After adjustment for patient preoperative characteristics and the complexity of the operations, there was no significant difference in the adjusted morbidity rates in three of the seven specialties and four of the eight common operations. Significant differences in postoperative morbidity rates after adjustment for patient risk and the complexity of the operation were found in general surgery, orthopedics, urology, and vascular surgery and in infrainguinal vascular reconstruction, partial colectomy, and open and laparoscopic cholecystectomy. Persistent differences in morbidity rates between the teaching and nonteaching hospitals after risk adjustment may result from several factors. First, it could be due to limitations of the NSQIP risk adjustment models for morbidity. The predictive validity of the morbidity risk adjustment models in this study was consistently lower than the predictive validity of the morbidity risk adjustment models for mortality. Although most the mortality models' c-indexes were 0.89 or more in surgical specialties and 0.77 or more in the prevalent operations, most of the morbidity c-indexes were 0.73 or less in the surgical specialties and 0.66 or less in the prevalent operations. In some specialties and procedures, the low c-indexes in morbidity (vascular surgery 0.63, lobectomy/pneumonectomy 0.63, carotid endarterectomy 0.59, infrainguinal vascular reconstruction 0.59) indicate poor predictive validity. The risk adjustment models for morbidity may not be as robust as the mortality models in accounting for all the differences in patient risk factors important for assessing differences in morbidity rates between teaching and nonteaching hospitals. The most frequent complications in most of the specialties and prevalent operations were wound morbidity, predominantly superficial and deep wound infections (see Table 6). The increased complexity of the operations in the teaching hospitals, coupled with significantly longer operative times, could partially account for the increased incidence of wound complications observed in these hospitals.

Differences in the structure and process of care between teaching and nonteaching hospitals may also be a cause for the increased morbidity rates in teaching hospitals. Typically teaching hospitals are larger and more complex than nonteaching hospitals, with more complicated systems of coordinating and managing care. Turnover among residents, fellows, medical students, and surgical attendings is high as all staff rotate through the affiliated teaching institutions. These more complex systems of care and constantly changing staff may create conditions in which

complications, particularly infections, occur more frequently. Hence, the increased morbidity rate observed in certain specialties and operations in teaching hospitals in this study might be due to the characteristics of the hospitals per se, and not necessarily because the operations in these hospitals were performed by residents.

Our findings are in part similar to those of other investigators. In a prospective VA cooperative study on valvular heart disease, there was no difference in risk -adjusted rates of morbidity and mortality between patients operated on for valve replacement by a resident under the supervision of attending surgeon or by an attending surgeon. [31] Several studies of outcomes using administrative data show lower mortality rates among patients cared for in large urban teaching programs than in community teaching and nonteaching community hospitals, but costs and length of stay are higher. [2] [3] [2] Most of these studies analyze large administrative databases and are limited to medical conditions such as stroke, pneumonia, congestive heart failure, and ischemic heart disease. Limited clinical information is available for risk adjustment with important clinical variables, and administrative data are subject to coding and other biases. Studies of the impact of teaching programs on surgical outcomes are limited to single surgical practices and low -mortality risk operations such as carotid endarterectomy and transurethral prostatectomy. The most recent large administrative database study [7] of 10 medical conditions and 10 surgical operations concluded that Medicare patients cared for in not-for-profit teaching hospitals had significantly lower risk-adjusted mortality rates but longer lengths of stay compared with five other hospital types (not-for-profit, osteopathic, public, teaching public, and for-profit hospitals). Among Medicare patients undergoing prostatectomy, lower extremity fracture repairs, bowel resection, vascular surgery, hip replacement, hysterectomy, mastectomy, laminectomy, or endarterectomy, adjusted 30 - day mortality rates were significantly lower in almost all comparisons, whereas the risk-adjusted length of stay was invariably longer. None of these studies reported on postoperative morbidity.

The results of this study in the VA may not be generalizable to nonveteran populations because of the unique features of the VA healthcare system and the population for which we care, mostly economically disadvantaged and disabled men. Risk adjustment models for individual operations and surgical specialties may be limited by the lack of more detailed risk adjusters and operation-specific postoperative complications that might change the results of the morbidity analysis. Finally, despite ongoing data verification and training of nurse reviewers, ascertainment bias in identifying postoperative adverse events may lead to under- or overreporting.

Our findings present an opportunity to identify systemic reasons for the higher morbidity rates observed in certain specialties and after certain operations in teaching hospitals. Ongoing investigations of postoperative adverse events in the VA [34] and improvement programs in surgery in each VAMC, as well as better surgical and anesthetic techniques, have led to a 45% decline in postoperative morbidity in the NSQIP between 1991 and 2000 (NSQIP, unpublished data).

Despite a cutback in surgical training positions in the VA to increase the availability of primary care physicians, the VA surgical postgraduate program still provides an important source of surgical training and staffing. From the VA's perspective, we have continued evidence of the good quality of care rendered in surgery to a vulnerable population and a vital role in educating the surgeons of the future. Opportunities in the near future to compare the outcomes and efficiency of care between the VA teaching and nonteaching programs and those in nonfederal hospitals will allow further comparison of the severity of illness, surgical complexity, process of care, and outcomes in surgical care and surgical education.

In conclusion, teaching hospitals perform 80% of major surgery in the VA. They attend to a sicker and more complex population of veterans than non-teaching hospitals, with a comparable risk - adjusted 30-day postoperative mortality rate. In four of seven surgical specialties and four of eight prevalent operations, the 30-day postoperative morbidity rate is higher in teaching than in nonteaching hospitals, even after risk adjustment. The decreased c -indexes of the morbidity versus the mortality models may reflect limitations in the risk adjustment methodology for morbidity. It may also reflect suboptimal processes and structures of care unique to teaching hospitals. These findings will prompt the NSQIP to focus further on the assessment and improvement of the structure and processes of care in the teaching hospitals, which are vital to the VA's mission to deliver quality healthcare to all veterans.

DISCUSSION

Dr. Jonathan L. Meakins (Montreal, Quebec, Canada): This is a very complex manuscript which I have had the privilege to read and which I have to admit I still don't completely understand, having gone through it several times.

The key findings that are herein relate to those associated with increased morbidity in four of the specialties and four of eight of the common procedures.

There is within the VA data collection system a method of identifying who was the primary surgeon. In 90% of the cases it was a resident in the teaching hospitals and 50% of the time a resident who was an R-5 or greater. This is supported in fact by the complexity of the cases, by the clearly longer time it took for all of these operations to be done, which tends to indicate that the residents did the majority of the cases. The data supports clearly the difference between the teaching and the non -teaching hospitals.

In your past studies of risk adjusted outcomes which you referred to, you have noted the importance of systems and processes of care which are supported by other papers presented to this organization where hospital experience and volume rather than the particular surgeon who operates in complex situations impact on the excellence of outcome. How can we take these raw data and turn them into a quality improvement exercise, which is clearly what the public wishes to have us do?

Let me separate surgical care rather arbitrarily into three component parts. There is the preoperative component part which has to do with the decision of who to operate upon and which operation to do, then the operative issues, and then the postoperative care issues, all of which are related to different levels and intensity of supervision.

Indeed, the morbidity, which has not been presented, can be broken out into two components. One is the operative related issues, which are wound morbidity, clearly higher in the teaching hospitals, and then the medical morbidity such as pneumonia, myocardial infarcts and so on.

I suspect the wound morbidity are primarily related to the attending asking the resident would they mind closing while the attendant goes off to do something else. I would ask if you have any data to show that this particular complication can be reduced by the changing process of the care and the organization of care in the operating room. The medical morbidity, on the other hand, occur in the third component of supervision, but perhaps can relate also to preoperative decisions related to operating.

How then do you intend to factor out the variables related to supervision that may have a role in the apparent differences in morbidity in the teaching institutions? The VA has rules about attending staff presence in the operating room, can they apply to preoperative decision making, the organization of the operation, and indeed postoperative management?

Presenter Dr. Shukri F. Khuri (West Roxbury, Massachusetts): The NSQIP is primarily a quality improvement program. The advantage of it is that it provides reliable data collected by dedicated nurses at each site. It is through the provision of reliable risk -adjusted outcomes that we have come in the VA to be able to identify weaknesses in processes of care, as Dr. Meakins has alluded to.

How can we turn the NSQIP data collection process into quality improvement initiatives? Surgeons in the VA continuously use the comparative data provided to them by the NSQIP to identify and improve suboptimal care within their respective institutions locally. (Reference #34 in our manuscript describes specifically one such initiative). The study, which I presented to you this afternoon, showed that risk-adjusted morbidity rates in certain subspecialties and certain operations were higher in teaching than non -teaching hospitals, suggesting inferior processes of care in teaching hospitals. These data will now trigger a focused review, which will be conducted by the NSQIP, to identify the causes for the higher morbidity rates in teaching hospitals and to ascertain to what degree is the presence of a teaching program in these hospitals an etiologic factor. Considering that most of the morbidity was in the form of postoperative infections, it is quite possible that the size and other characteristics of the teaching hospitals, not necessarily the presence of residents or the degree of attending supervision, were the factors primarily responsible for the increased morbidity rates. A focused review which identifies these factors and the processes which need to be implemented to address them, will ultimately bring about quality improvement, particularly since the VA has a quantifiable intricate system for assessing the adequacy of resident supervision.

Dr. Alden H. Harken (Denver, Colorado): Dr. Khuri, thank you very much for letting me review your manuscript. I think we are all tremendously indebted to you for developing these kinds of quality assurance programs in identifying some of the six competencies that they have identified. This kind of patient based learning is very important, and if you don't do it, as multiple people already discussed in this meeting, some green-visored weasel in a regulatory agency in Washington is going to do it for us. So I think we all have to do what you are doing.

In that regard, as I look through the kinds of statistical problems we have as surgeons in doing clinical investigation, which is exceedingly difficult to do, the typical one is the type 2 beta error. We don't have enough patients in a study to make the observation that we are making.

And I hold in my mind that in order to pick up a 20% difference, which is a big one, at the 95% confidence level, you have got to have 100 patients in both groups. And I looked through the 32 manuscripts at this meeting, and half of them have more than 100 patients in both groups. So the quality of the studies being presented here is really special.

On the other end of that spectrum, however, with 690,000 patients in your study, you are at risk of the type 1 or alpha error. And I looked at your general surgical population, the mean age in the teaching hospitals was 60.8 years and the mean age in non-teaching hospitals was 61.8 years, one year difference. And although I am approaching that with perilous velocity, I am hoping I don't fall off the cliff. But that one year was statistically different at the .0001 level.

Can you comment on the risks or the differences between statistical significance and clinical relevance and what is the risk of presenting these kinds of observations as statistically significant to a lay or regulatory audience?

Dr. Shukri F. Khuri: Thank you, Dr. Harken. You bring up a very important point. When we first started looking at data generated by the VA National Surgical Risk Study, the forerunner of the NSQIP, we were frequently struck by statistically significant differences, generated by comparing very large numbers of patients, which did not make clinical sense. One of the strengths of the NSQIP is that it is a collaborative effort between surgeons, health policy researchers and statisticians. Our colleagues have helped us devise analytic methods that void alpha (Type I statistical) errors, which might result from univariate analyses comparing large patient populations. In this study, for example, we employed sophisticated two-level hierarchical modeling, in addition to logistic regression, so as to account separately for the severity of illness of the patient, and the characteristics of the hospital. While it is likely that an alpha error may be committed in a large univariate analyses, it is most unlikely to encounter such an error after sophisticated multivariate analyses.

Dr. Carlos Pellegrini (Seattle, Washington): Like the previous discussants, I compliment you on working so hard on the improvement program in the VA and in creating a database that allows us to predict outcome. I have some specific questions.

First, several recent publications that examined the issue of teaching versus non-teaching hospitals have found that teaching hospitals do better in terms of mortality and morbidity than non-teaching hospitals in urban areas. If one thinks about it, the teaching institutions are a repository of unique specialty, intellectual and infrastructure resources, and so on. Your study found increased morbidity in teaching hospitals. Can you identify anything that the VA might be doing different than other teaching institutions?

Second, I would like you to explain the definition of a teaching hospital. You define a teaching hospital as one with a dean's committee and with at least a residency program. However, in reading the data in your manuscript, somewhere between 1.5 and 15% of the operations done in a non-teaching hospital have a resident listed as a primary surgeon. So this calls into question the definition of a teaching hospital. If one thinks for a moment that the presence of the resident may be an important variable in the outcome, as Dr. Meakins so eloquently just pointed out, wouldn't it be important to also run this analysis in such a way as to convert into "teaching hospitals" any operation performed by a resident as the primary surgeon? Have you done that?

Thirdly, the study did not control for surgical volume of the hospital or the surgeon's case experience. Both of those variables have been shown to be linked to outcome independent of teaching status. Can you do that with this database?

Fourth, is it possible to measure morbidity in a manner that is specific to a procedure rather than in general? For example, have you analyzed common bile duct injury during cholecystectomy, or dehiscence in the case of colectomy, et cetera, rather than more general types of morbidity?

Lastly, just a philosophical comment. And that is that when all the statistics are done and examined, I think that we need to go back to basic judgment and realize that in our profession, like all others, it is always going to be a very difficult challenge to train individuals on the job and obtain the same results that could potentially be obtained if only fully trained surgeons perform the task. I think it is our collective responsibility to organize systems that at least minimize the negative effect associated with the lack of knowledge or experience among residents. Assuring, for example, that residents have sufficient rest periods, maximizing the time spent in a given service, minimizing rotations, and developing new systems of learning that allow residents to achieve a certain level of performance before taking the full responsibility for patients I think is going to be vital to keeping a system like this.

Dr. Shukri F. Khuri: Thank you, Dr. Pellegrini, for your insightful comments. In the wake of the morbidity data, it is important not to underestimate the significance of the risk-adjusted mortality rates in this study. The NSQIP has validated and uses the risk-adjusted 30-day mortality rate as a comparative measure of the quality of surgical care. The fact that we could not elicit a difference in the risk-adjusted mortality rates between teaching and non-teaching hospitals, despite a marked increase in the severity of illness of patients in teaching hospitals indicates that the quality of care is comparable in both types of hospitals.

The designation of a teaching hospital and the analyses were based on the surgical specialty. At least one resident in a specialty was necessary to designate the hospital as a teaching hospital in that specialty. The reason why up to 15% of the operations in a certain specialty were performed by residents and yet that specialty was labeled "non-teaching" was because the residents did not belong to that particular specialty. For example, when a general surgery resident was listed on a thoracic operation. Originally, we performed the analyses as you suggested, separating the two groups on the basis of whether or not the operation had been performed by a resident, irrespective of the specialty. The mortality results were no different from those reported in this study.

We do not collect operation-specific information in the NSQIP database. The generic preoperative risk factors, which we collect, are enough to generate highly predictive models in mortality. The predictive validity of our morbidity models will be enhanced if condition-specific variables are collected.

The findings in our study that VA teaching hospitals operate on higher risk patients than non-teaching hospitals, with comparable mortality rates, speak for what the literature has already suggested: after accounting for the difference in risk and complexity, the mortality rate in teaching hospitals is probably better than in non-teaching hospitals. There is little in the literature on comparative morbidity rates between the two types of hospitals. Our findings that in some specialties and certain operations these rates are higher in teaching hospitals, even after risk adjustment, is probably a reflection of the limitations of our risk adjustment models which are much less robust for morbidity than mortality. Nevertheless, they do call for a focused review which should answer most of the questions you raised, particularly issues related to the impact of residency supervision and the degree of attending involvement within the continuum of surgical care.

Dr. Lawrence W. Way (San Francisco, California): It is ironic that we are discussing the quality of care in teaching hospitals.

At the end of World War II, a group of prominent physicians and surgeons was asked for recommendations on how best to manage the care in VA hospitals of the large number of new veterans. Creating University affiliations was thought to be important, because the government was felt to have an obligation to participate directly in medical education, and also because affiliating with Universities was considered to be the best way to ensure that veterans would receive high quality care.

References

1. Blumenthal D, Meyer GS. The future of the academic medical center under health care reform. N Engl J Med 1993; 329: 1812-1814.   Citation

2. Iglehart JK. Rapid changes for academic medical centers. N Engl J Med 1994; 331: 1391-1395.   Citation

3. Epstein AM. US teaching hospitals in the evolving health care system. JAMA 1995; 273: 1203-1207.   Citation

4. Iglehart JK. Rapid changes for academic medical centers, II. N Engl J Med 1994; 332: 407-411.

5. Kassirer JP. Academic medical centers under siege. N Engl J Med 1994; 331: 1370-1371.   Citation

6. Whitcomb ME, Cleverly WO. Financial performance of academic medical center hospitals. Acad Med 1993; 68: 729-731.   Abstract

7. Rosenthal GE, Harper DL, Quinn LM, et al. Severity-adjusted mortality and length of stay in teaching and nonteaching hospitals. Results of a regional study. JAMA 1997; 278: 485-490.   Abstract

8. Keeler EB, Rubenstein LV, Kahn KL, et al. Hospital characteristics and quality of care. JAMA 1992; 268: 1709-1714.   Abstract

9. Taylor DH Jr, Whellan DJ, Sloan FA. Effects of admission to a teaching hospital on the cost and quality of care for Medicare beneficiaries. N Engl J Med 1999; 340: 293-299.   Abstract

10. Hartz AJ, Krakauer H, Kuhn EM, et al. Hospital characteristics and mortality rates. N Engl J Med 1989; 321: 1720-1725.   Abstract

11. Iezzoni LI, Shwartz M, Moskowitz MA, et al. Illness severity and costs of admissions at teaching and nonteaching hospitals. JAMA. 1990; 264: 1426-1431.   Abstract

12. Japsen B. Teaching hospitals face hard lessons. Mod Healthcare 1994; 7: 36-8.

13. Sloan FA, Feldman RD, Steinwald B. Effects of teaching hospital costs. J Health Econ 1983; 2: 1-28.   Abstract

14. Sloan FA, Valvona J. High cost of teaching hospitals. Health Aff (Millwood) 1986; 5 (3): 68-85.   Citation

15. Zimmerman JE, Shortell SM, Knaus WA, et al. Value and cost of teaching hospitals: a prospective, multicenter, inception cohort study. Crit Care Med 1993; 21: 1432-1442.   Abstract

16. Medicare and the American health care system: report to the Congress, June 1,1997. Washington DC: Prospective Payment Association Commission, 1997.

17. Yuan Z, Cooper GS, Einstadter D, et al. The association between hospital type and mortality and length of stay: a study of 16.9 million hospitalized Medicare beneficiaries. Med Care 2000; 38: 231-245.   Abstract

18. Estey EP, Mador DR, McPhee MS. A review of 1486 transurethral resections of the prostate in a teaching hospital. Can J Surg 1993; 36: 37-40.   Abstract

19. Bottsford JH, Blake R, Small T, et al. Carotid endarterectomy: the outcome is not adversely affected by surgical residents. J SC Med Assoc 1997; 93: 243-247.

20. Evans SM, Adam DJ, Murie JA, et al. Training in abdominal aortic aneurysm (AAA) repair: 1987-1997. Eur J Vasc Endovasc Surg 1999; 18: 430-453.   Abstract

21. Kizer KW. The "New VA": a national laboratory for health care quality management. Am J Med Quality, 1999. :

22. Khuri SF, Daley J, Henderson WG, et al. The Department of Veteran's Affairs NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. **Ann Surg** 1998; 228: 491-507.   Full Text

23. Khuri SF, Daley J, Henderson W, et al. The National Veterans Administration Surgical Risk Study: risk adjustment for the comparative assessment of the quality of surgical care. J Am Coll Surg 1995; 180: 519-531.   Abstract

24. Khuri SF, Daley J, Henderson W, et al. Risk adjustment of the postoperative mortality rate for the comparative assessment of the quality of surgical care: Results of the National VA Surgical Risk Study. J Am Coll Surg 1997; 185: 315-327.   Abstract

25. Daley J, Khuri SF, Henderson W, et al. Risk adjustment of the postoperative morbidity rate for the comparative assessment of the quality of surgical care: Results of the National VA Surgical Risk Study. J Am Coll Surg 1997; 185: 328-340.   Abstract

26. Buck S. A method of estimation of missing values in multivariate data suitable for use with electronic computer. J Royal Stat Soc 1960; 22: 302-307.

27. Bryk AS, Raudenbush SW. Hierarchical linear models: applications and data analysis methods. Newbury Park, CA: Sage; 1992.

28. Normand SL, Glickman ME, Gatsonis CA. Statistical methods for profiling providers of medical care: issues and applications. J Am Stat Assoc 1997; 92: 803-814.

29. Wong GJ, Mason WM. The hierarchical logistic regression model for multivariate analysis. J Am Stat Assoc 1985; 82: 613-618.

30. Kizer KW, Pane GA. The "New VA": delivering health care value through integrated service networks. Ann Emerg Med 1997; 30: 804-807.   Full Text

31. Daley J, Forbes MF, Young G, et al. Validating risk-adjusted surgical outcomes: site visit assessment of process and structure. J Am Coll Surg 1997; 185: 341-351.   Abstract

32. Gibbs JO, Khuri SF, Clark K, et al. Validating risk-adjusted surgical outcomes: chart review of process of care. Int J Health Quality (in press).

33. Sethi GK, Hammermeister KE, Oprian C, et al. Impact of resident training on postoperative morbidity in patients undergoing single valve replacement. Department of Veterans Affairs Cooperative Study on Valvular Heart Disease. J Thorac Cardiovasc Surg 1991; 101: 1053-1059.   Abstract

34. Neumayer L, Mastin M, Vanderhoof L, et al. Using the Veterans Administration National Surgical Quality Improvement Program to improve patient outcomes. J Surg Res 2000; 88: 58-61.   Abstract